A Decision Model for Retail Location in Indian Cities

Synopsis

PhD Dissertation in Urban Planning

Submitted by

Joydeep Dutta

Under the supervision of

Prof. Dr. Uttam Banerjee

Department of Architecture & Regional Planning

Indian Institute of Technology Kharagpur 721302

August 2007

A Decision Model for Retail Location in Indian Cities

Synopsis for PhD Dissertation in Urban Planning

Submitted by: **Joydeep Dutta**

Under the supervision of: Prof. Dr. Uttam Banerjee

Department of Architecture & Regional Planning,

Indian Institute of Technology Kharagpur 721302, India

August 2007

This study investigates the measure of attractiveness of retail stores and formulates a decision model based on Multiplicative Competitive Interaction (MCI) to predict viable urban retail locations. It identifies and estimates the parameters that determine relative utilities of urban locations and reflect potential market share. The empirical data used to calibrate and validate the model are based on a case study of organized sector clothing stores in the city of Kolkata, India. The study provides a simple decision model to urban planners, as well as the business and management science communities, for making informed choices regarding intra-urban retail locations, and discusses notable findings regarding prediction models, consumer behaviour and urban planning implications.

---

1. Introduction

Retail is the new global phenomenon and India is a large market that is quickly waking up to it. Organized retailing as well as the traditional unorganized retail sector is set to gain immensely in the near future in terms of revenue and employment. The current globalization of attitudes and consumer preferences in India, make retailing the most lucrative commercial prospect in all major cities. Retailing as a science and broad area of research is not new to the developed world, presently dominated by business managers and management scientists, when actually it is urban planners who are better conversant to do proper justice to it, especially in the domain of retail location analysis.

Traditionally, small independent shops have dominated retailing in the developing world. India has sometimes been referred to as 'a nation of shopkeepers.' These small shop owners sell products of a particular category, operate from fixed small-to-medium format outlets and mostly restrict

themselves to specific geographic locations of their convenience. Thus, decision making for these small retailers presently does not entail a thorough knowledge of the various factors that influence marketing management and affect market share total sales volume.

The Indian consumer is fast evolving. Changing consumer preferences, increased personal incomes, greater spending power and a proliferation of branded goods, have changed the urban spending scenario drastically over the last decade. Large companies engaged in business (and more specifically retailing) have to deal with a complex set of market forces and thus have to adopt retailing strategies in a more informed and well thought-out manner. Extensive market research and consumer behaviour studies are needed to make this possible. In the recent past, corporate giants and manufacturers have sometimes engaged in such database building and research to help them in their retail-oriented decision-making. But these studies have been conducted either privately or by market study groups in a piecemeal manner. The data collected by them or the insights gained by them have never been shared, thus being of little or no real use for further research and development.

1.1 Need for the study

With the liberalization of Indian economic policies and the increase in expendable income of the urban population of India, there is tremendous growth in the organized retail sector since the turn of the century. Of all the strategic decisions of the retailer (price, product mix, location and scale of operations), location remains the most inflexible, irreversible and thus critical decision. The research domain of location studies pertaining to retail is vast and numerous decision making models have been used in the USA and Europe. In developing economies like India, retail location decisions are done in a very ad-hoc semi-informed manner, based on gut-feeling of the retailer (Biyani, 2001). There is presently a definite lacuna in the research and development of location decision modelling for Indian cities based on reliable urban and retail competition data.

Most Indian cities are historic and organic in growth. It is possible to identify the history of retail growth in a region, but it is not easy to ascertain the demand patterns that emerge in retail outlet locations. In the current context of globalization of markets, many international manufacturing and retail establishments identify Indian cities as large potential markets. But they find it almost impossible to discover a specific methodology to locate their outlets and ascertain the scale and type of operations for maximum economic benefit in a particular geographic region. The complex and multi-disciplinary domain of decision-making for retailers, especially retail location needs to be examined critically by urban planners and geographers in developing countries like India. The general approach adopted by urban planners is indeed the right mix of spatial and socio-economic expertise required for the emerging field of retail planning. Integrated urban models integrating land use models and transportation models have been attempted before (Putman, 1983), but they need to be modified and made applicable in situations of data inadequacy. Hence the present study.

There is a serious lack of integrated databases of existing retail outlets and marketing agglomerations in Indian cities and towns. The few studies that exist are conducted by market research organisations and relate only to specific retail categories (ORG MARG, 1995). The few consumer studies in India are conducted through newspaper readership surveys and relate to socio-economic stratification (NCAER 2001, IMRB 2001). International comparisons of retail trends are often the easiest way to assess the direction of consumer demand and choice behaviour. But there needs to be a sound critical analysis of the buying power and shopping behaviour of the Indian consumer, which are much divorced from the trends we find in developed countries. The real challenge lies in judiciously overlaying these socio-economic considerations over the existing spatial distribution of the urban opportunities and retail developments of a city.

Urban planners are familiar with the existing spatial distribution of economic activities within the city, carrying capacities of the road network, infrastructure provisions and also the dynamics of land-use classifications over time. An effective decision model for intra-urban retail locations will help prospective retailers to make informed choices for locating their outlets, help economists and management scientists to forecast future sales and market share trends, and finally help urban planners to frame urban spatio-economic policies and land-use allocations over space and time.

1.2 Broad goal & objectives

The broad goal of the study is to devise an effective decision model for predicting suitable retail locations in a given urban network, by determining relative attractiveness of intra-urban nodes. Towards this goal, the major objectives are:

- To create a database of existing markets within the urban network, incorporating geographical and socio-economic data from the supply side (case study of organized clothing in Kolkata)
- To determine the attractiveness potential for all possible store locations in the network based on market share determinants
- To predict a set of best-suited possible locations for potential retailers wishing to open outlets in the given network.

The study methodology is presented as an overview in Fig.1.

1.4 Scope of the study

The basic decision framework used in this study is the retail gravity model which predicts market share of existing retail locations, based on relative utility of an intra-urban node vis-à-vis all existing intra-urban nodes within that city (Huff, 1963). The proposed model formulation is based on Luce's Choice Axiom using Multiplicative Competitive Interaction (MCI), where the product of multiple independent predictors is used to predict relative attractiveness of locations.

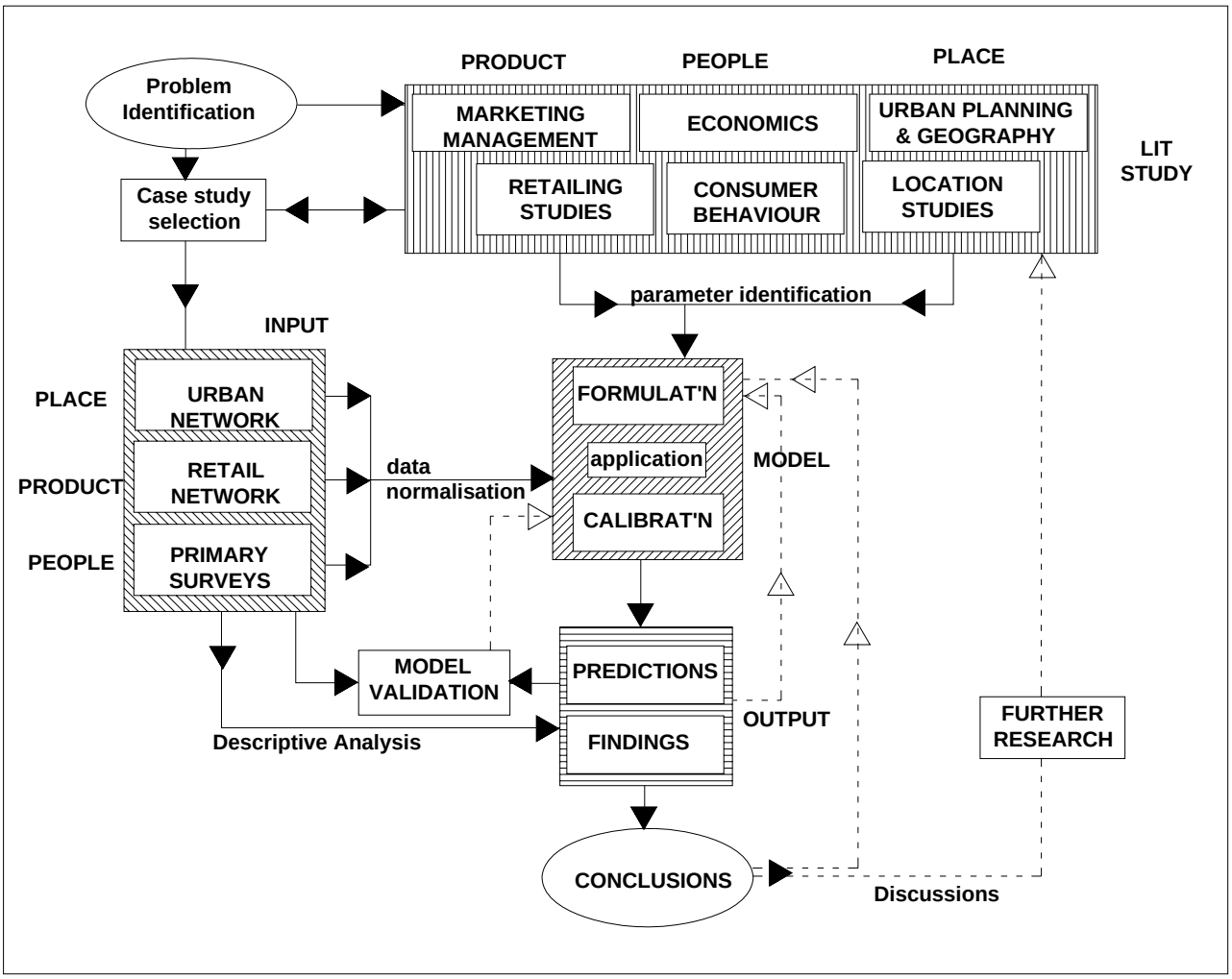


Fig.1 Overview of Study Methodology

Location attractiveness (LA) and Store attractiveness (SA) are the two independent variables used to predict the market share (dependent variable), as observed from store footfall in the existing scenario. The LA of a node is estimated by four determinant parameters – demography, travel time, competition and urbanization. SA measures are store format, floor space, number of employees, total investment, etc. but the estimation for SA variables is beyond the scope of the work.

MCI has been employed to determine the relative importance of the location attractiveness parameters, by estimating their respective exponents and calibrating the proposed model against observed empirical data. The resultant equation is then employed to predict the expected market shares of candidate intra-urban nodes, ranking the most attractive retail locations within the city. The model is validated by using Mean Absolute Percentage Error (MAPE) method of comparing predicted values with observed market share values that were not initially used to run the model. Sensitivity analysis of the predicted attractiveness results is conducted across the three different store formats possible – large, medium and small.

Three C programming computer codes are used in the study to:

(1) Calculate minimum travel time matrix between all nodes using Floyd's algorithm

(2) Calibrate the exponents of the parameters by ordinary least squares method (OLS), using log-linear transformation (Nakanishi & Cooper, 1974)

(3) Apply the proposed model for ranking the candidate nodes by highest attractiveness

Finally, the study uses three different mathematical formulations (multiplicative, additive and exponential) to compare the interaction between the two attractiveness measures LA and SA. This synopsis follows the pattern of chapterization and numbering of the final thesis.

2. Literature Review

The literature review for the study encompasses three major fields of research interest – marketing management, economics and urban geography. Here, the domain of review is limited to the organized retail scenario in India. Location theories and prediction techniques, both classical and modern, are reviewed briefly in light of their applicability to the present study. Only the directly relevant theories are mentioned here, i.e. spatial interaction theories, consumer beahviour theories and information-based decision support systems (DSS). The literature review attempts to converge on the determinant factors that influence retail location and will be used in the study.

2.1 Organized Retailing in India

Retail opportunities are presently booming in India with increased globalization and the rise in disposable income of the urban consumer. The retail industry presently accounts for about 11% of the GDP and is one of the largest employers of urban India. Traditionally in India, retailing was dominated by small one-owner shops and corner stores ('kirana' shops). There are about 6 million retail establishments in India, accounting for about 6.7% of the total employment (Census, 2001).

India ranks first on the Global Retail Development Index (a study of retail investment attractiveness among 30 emerging markets), up from sixth rank in 2004 (ATKearney, 2007). India's retail sector is estimated worth US$ 350-380 billion, of which only about 6% was in the organized sector in 2006. Organized retailing refers to improved store formats with large floor space and wide product mixes, that provide quality products at fixed prices with improved customer service. In India, this sector can presently be categorized under the categories: (i) Food retailers, (ii) Health and beauty products, (iii) Clothing and footwear, (iv) Home Furniture and household goods, (v) Durable goods, (vi) Leisure and personal goods.

Organized retailing in India has started only over the last decade and has been concentrated mainly to the metros. However, it will continue to grow rapidly, especially through increased levels of penetration in larger towns and also begin to spread to smaller cities and B class towns (CII, 2004). This is being reflected in the rapid growth in sales of malls, supermarkets, departmental stores and hypermarkets – all varying business versions of large-store formats. Organized retailing would

increase the efficiency and productivity of the entire gamut of economic activities of the nation, achieve higher GDP growth and provide huge employment at all levels. More than 25 million sq. ft. of retail space is being developed and will be completed by 2007, and organized retailing in India will capture about 15 – 20% of the market share by 2010 (Jones Long LaSalle 2007).

2.2 Retail Location Studies

Location studies have been carried out the world over for more than a century now. Classical concepts such as economic theories, central place theories, land value theories, and spatial interaction theories dominated till the 1950's. The next twenty years saw more sophisticated and complex models with the advent of computing, but these neo-classical studies have sometimes been criticized as being too autistic, far removed from reality. Thereafter, the role of social scientists was acknowledged in decision theory and consumer studies overshadowed hard-core spatial analysis for location. Modern day information systems and management science methods are rigorous and detailed, but they are effective only when large-scale real time-series data is available. The cutting-edge research in this field is attempting to simulate the complexities of human behaviour and understanding by artificial intelligence.

Efforts at literature reviews on location studies have proposed different morphologies for summarizing retail location systems (Scott 1970, Revelle et al 1970, Kivell and Shaw 1980). Built on these, the following table (Fig. 2) provides a chronological morphology for various retail location theories over time (Dutta & Banerjee, 2002). Only spatial interaction theories, consumer behaviour models and information-based DSS are discussed here, as they directly relate to the study.

THEORY	**Classical**	**Neo-classical**	**Contemporary**
Economic theories	Population estimates Demand assessment	Market potential Harris 1954	Marketing databases Simulation and gaming
Central Place theories	Hotelling 1929 Christaller 1933	Lösch 1954 Berry-Garrison 1958 Retail agglomeration	G.I.S. Socio economic data Census data
Land value theories	vön Thunen Bid rent curve	Theory of urban rents Scott 1970	Real estate forecasting
Spatial Interaction theories	Reilley's Law 1931 Probabilistic Utility	Huff 1963 Lakshmanan-Hansen 1965 Harris 1964 Lowry 1964	Network algorithms Heuristics
Iterative and heuristic models	Weber 1929 Pick, Varignon 1929	Mathematical formulations Operations research	Advanced micro-location techniques
Consumer behaviour studies		Fine 1954 Cognitive learning cycle Aggregate behaviour	Choice modelling: Observed behaviour Intended behaviour
Information-based decision support systems		Decision under uncertainty Real time-series data management science methods	Decision Support and Expert systems Artificial Intelligence

Fig. 2: Morphology for Retail Location Theories

2.3 Spatial interaction theories

This is the most widely used and accepted retail location theory and has its origin in the age-old gravitation concept of human interaction. It postulates that 'an attracting force of interaction between two areas of human activity is created by the population masses of the two areas, and a friction against interaction is caused by the intervening space over which the interaction must take place' (Carrothers, 1956). The Law of Retail Gravitation (Reilley, 1931) states that the force of retail attraction between two cities is directly proportional to their respective populations and inversely proportional to the square of the distances between them. Reilley makes two major assumptions – that two competing trade areas are equally accessible along a road and that the merchants in the two areas are equally effective. The concept of determining the equilibrium point (from where a consumer will be equally attracted to two competing markets) is then employed.

The concept of probabilistic utility is central to the understanding of equilibrium analysis and spatial interaction models. The probability of shopping at outlet *j* is the utility of shop *j* divided by the total utility of all shops in a market. It takes the form $P_j = U_j / \sum P_j$. Several studies have later modified this concept and introduced parameters relating to consumer behaviour to sophisticate it further. The first significant attempt to bridge the gravity formulation and consumer behaviour was Huff's Law of Shopper Attraction (Huff, 1963). He expressed the probability of a consumer to travel to a particular store, using two variables – store-size (floor space) and travel time and estimated different travel-time exponents for different retail types.

Based on this theory, many location models were formulated and field studies conducted in the 1960's (Lakshmanan & Hansen, Haydock Proposal, South Bedfordhire, etc.) using real time-series data on sales and consumer expenditure. The Penn-Jersey Transport study uses behavioural parameters that govern the attenuation of trips over distance and over competing facilities (Harris, 1964). The Pittsburgh model, used specifically for the retail sector, indicates that each retail type has a characteristic production function (expressed by minimum size constraints) and specifies the number of clients and the amount of floor space required to support one employee (Lowry, 1964).

2.4 Consumer behaviour theories

The perception of consumers and shopping behaviour emerged as important factors in determining retail strategies (Fine, 1954). This necessitated the quantification of subjective human phenomena, giving rise to psychometric analysis by a three- or five-point scale. The role of the questionnaire as a detailed research tool emerged universally. The primary assumption of these theories is that consumer behaviour in the aggregate can be predicted on the basis of mathematical probability. Some prediction models assess market share and market penetration of a store, while others use computerized trade area mapping systems by identifying homogeneous market segments, wholly comprising of similar consumers with a unique pattern of purchasing, consumption, media exposure and financial behaviour (VISION, 1987).

Contemporary retail behaviour studies may be summarized under three different approaches – normative, cognitive and empirical. Modelling consumer choice behaviour may be broadly classified into two categories based on (a) Observed choice behaviour and (b) Intended choice behaviour (Borgers & Timmermans, 1991). Observed choice theories stem from either spatial interaction models (after Wilson, 1971) that examine origin-destination data and actual consumer flows from residential zones to retail zones, or from random utility models denoting shop potential or attractiveness. However, intended choice behaviour may adopt either the stated preference approach or the choice calibration approach. The stated preference is a heuristic approach that is very popular in current business practices. It constructs hypothetical choices, asks the consumer to rank them by preference, simulates this behaviour on an existing database, and finally suggests a mix and match of existing shopping characteristics. The choice calibration approach measures the choices on a psychometric scale instead of constructing hypothetical choices, by using random utility functions from experimental design data.

2.5 Information-based decision support systems

The other major development in recent years is the extensive use of information systems and other sophisticated computer-supported management science methods. The complexities of human behaviour and understanding, earlier considered as the sole prerogative of psychologists and social scientists, were acknowledged by management scientists and managers within their domain of quantitative and analytical decision making processes. Artificial intelligence techniques essentially try to combine retail expertise and network characteristics of a particular existing setting, thus culminating into advanced retail models used in management science.

Decision-making under economic and spatial uncertainty involve scientific means of dealing with real-life complex decisions, with the help of management information systems such as Decision Support Systems (DSS), Expert Systems and Knowledge Based Systems (Keen, 1978, Turban 1990). A DSS is defined as ‘an interactive computer based system which supports managers in making unstructured (or semi-structured) decisions’ (Sprague, 1974). All such systems basically have a common three-tier structure: Database, Model-base and Graphic User Interface (GUI).

The detailed time-series data is organised and managed by the database, the model-base contains the algorithms and heuristics to be applied on the data for the required analysis, and the GUI provides the graphic platform to interact with potential users of the DSS. Some notable contributions to the specific field of retail location modelling include: dynamic spatial interaction models that help replicate structural changes in retail history, simulation models to estimate radical changes in retail location (Satani, 1998), and spatial DSS for retail plan generation and impact assessment (Arentze, 2000).

2.6 Factors influencing retail location

Recent marketing management literature that is reviewed reveals that retail decisions are based on four major parameters, which together constitute the retailer's strategy mix: (a) Pricing (b) Product Mix (c) Scale of Operation (d) Store Location. The pricing, product mix and brand image go hand-in-hand and is directly influenced by market segmentation, based on the socio-economic status of the prospective targeted clientele. The scale of operations for the retailer depends on the amount of investment he is willing to make, thus determining store format, size of store and number of employees, as well as product mix and store inventory. Finally, store location proves to be the least flexible and hence the most critical decision to make for the retailer. It involves high investments, tricky legal implications and is dependent on availability of space and the existing competition in the vicinity. New or prospective retailers thus attach great importance to the store location decision.

It is necessary to differentiate these retailing strategy decisions into spatial and non-spatial factors – those relating to "product" (store internalities), and those factors relating to the "place" (store externalities). Non-spatial factors, such as retail image and mix, significantly affect non-anchor retail patronage (Mejia & Benjamin, 2002) and are thus indirectly related to location decisions. **Store internalities** are pricing, store area, level of service, number of employees, store format, promotions, product mix, stocks, etc. **Store externalities** are site and location characteristics such as population, accessibility, rent and land value, amenities and infrastructure, location image and existing competition. Whereas the former decisions are the business manager's prerogative, the latter mentioned ones are the urban planner's prerogative.

Four distinct factors are identified for proper decision-making for store locations:

(a) **Customers** of the reference group, or target population – its size, density and growth trends, socio-economic characteristics such as buying power, image sensitivity
(b) **Costs** to set up shop – rents, land values, space availability, taxes and restrictions, etc.
(c) **Accessibility** of the location – road condition and image, existing infrastructure, access to public transportation, parking facilities
(d) **Competition** at the location – from competing land-uses, competition between different trades (indirect competition) and within a particular trade (direct competition).

3. Proposed methodology – rationale of the research

This section discusses the rationale of the present research and presents a schematic diagram of the modelling methodology, along with the basic inputs required (Fig 3).

Stage I

FORMULATION OF ATTRACTIVENESS MEASURES:

STORE Attractiveness **[SA]**
LOCATION Attractiveness **[LA]** (4 components for each node):

$$LA = \prod_{q=1}^{4} (D, T, C, U)$$ - The Multiplicative Competitive Interaction (MCI) form

D: Demography
T: Travel
C: Competition
U: Urbanization

Basic INPUTS
Input A: City Network

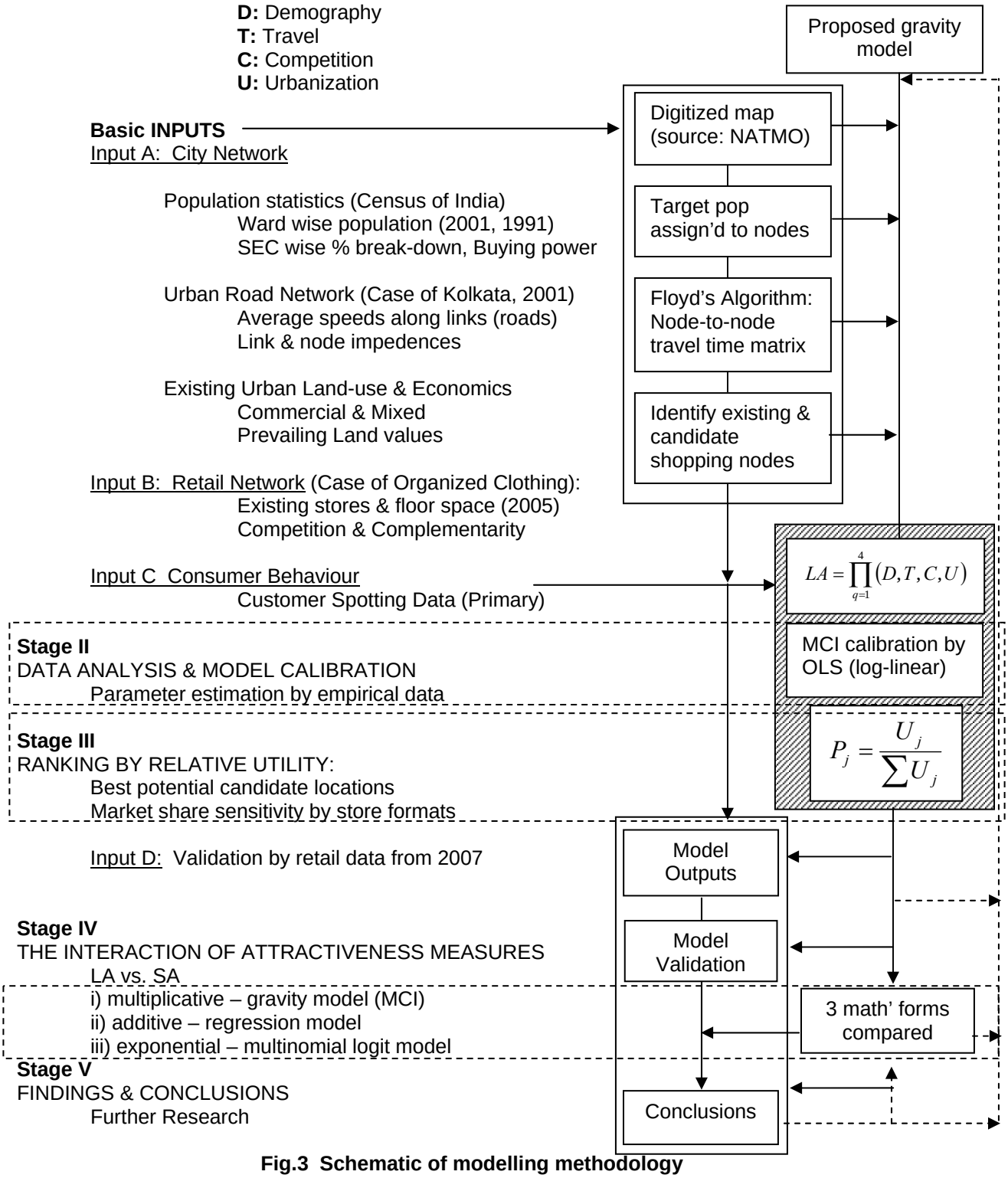
Population statistics (Census of India)
Ward wise population (2001, 1991)
SEC wise % break-down, Buying power

Urban Road Network (Case of Kolkata, 2001)
Average speeds along links (roads)
Link & node impedences

Existing Urban Land-use & Economics
Commercial & Mixed
Prevailing Land values

Input B: Retail Network (Case of Organized Clothing):
Existing stores & floor space (2005)
Competition & Complementarity

Input C Consumer Behaviour
Customer Spotting Data (Primary)


Proposed gravity model
Digitized map (source: NATMO)
Target pop assign'd to nodes
Floyd's Algorithm: Node-to-node travel time matrix
Identify existing & candidate shopping nodes
$LA = \prod_{q=1}^{4} (D, T, C, U)$
MCI calibration by OLS (log-linear)
$P_j = \frac{U_j}{\sum U_j}$
Model Outputs
Model Validation
3 math' forms compared
Conclusions

Stage II
DATA ANALYSIS & MODEL CALIBRATION
Parameter estimation by empirical data

Stage III
RANKING BY RELATIVE UTILITY:
Best potential candidate locations
Market share sensitivity by store formats

Input D: Validation by retail data from 2007

Stage IV
THE INTERACTION OF ATTRACTIVENESS MEASURES
LA vs. SA
i) multiplicative – gravity model (MCI)
ii) additive – regression model
iii) exponential – multinomial logit model

Stage V
FINDINGS & CONCLUSIONS
Further Research

Fig.3 Schematic of modelling methodology

3.1 Formulation of attractiveness measures

A basic contention of the study is that the location attractiveness (LA) of each given intra-urban node commands a unique relative market share depending on its geographic position and urban setting, which is unrelated and independent of the type or format of store located in it. Thus, LA and Store attractiveness (SA) are the two independent variables used to predict the market share (dependent variable), as observed from average store footfall (Fig.4). Location attractiveness (LA), i.e. the inherent potential of the place for retail, is a function of four distinct components, each of which will be considered an independent predictor variable:

(a) Demography (b) Travel Time (c) Competition and (d) Urbanization

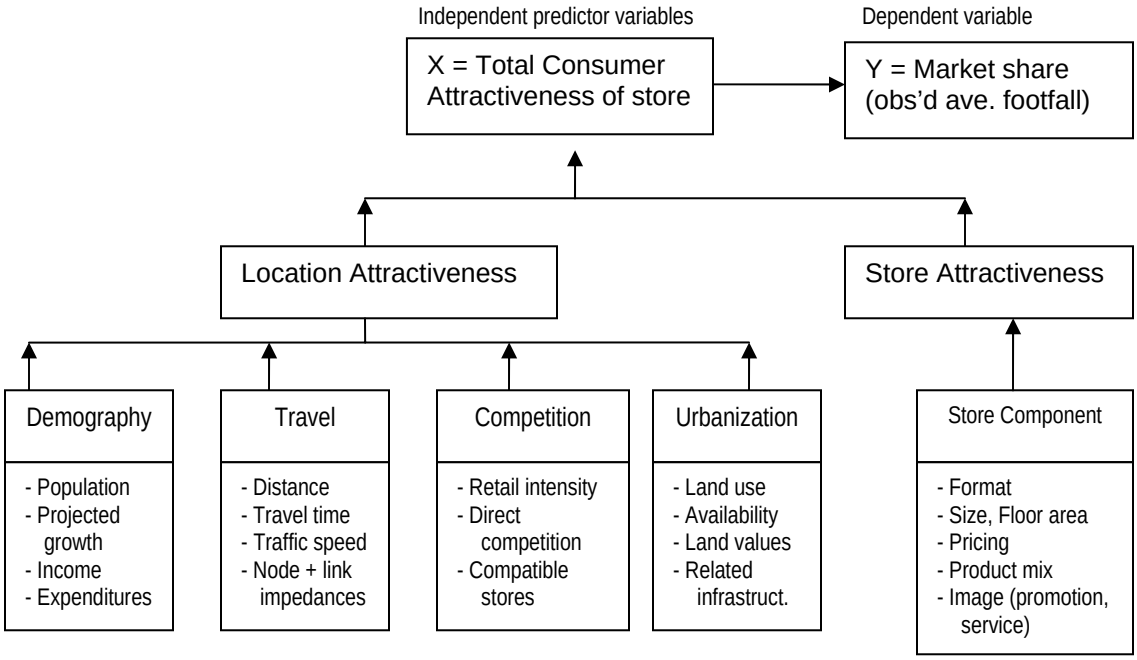


Fig. 4: Attractiveness measures of a store

Demography: The ultimate intention of this component is to determine the total consumer expenditure from a residential zone on the test retail commodity for the test time period. Time-series census data at a ward level is the first data requirement. The actual spatial distribution of population, measured as the projected number of resident households are assigned to every urban node on the city network. Market segmentation indicates the proportion of households that belong to a particular socio-economic category, pre-defined by the average monthly household income (MHI), likely to consume that particular product.

Travel Time: This component measures the consumer disincentive of making a shopping trip, providing a zone-to-zone matrix of travel impedances for a given city. This trip function has an inverse relationship with a consumer's actual trip-making probability. Traditionally, the impedance was measured in distance travelled, raised to a power of -2 (Reilley, 1931) and -1.33 (Lösch, 1954). Travel time is a better determinant of this disincentive than travel distances (Huff, 1963;

Lakshmanan, 1965), though the generalized travel cost may be assessed if origin-destination (O-D) data and transport mode-cost data are available. The customer-spotting technique is used to geometrically ascertain trading areas of stores – primary, secondary and tertiary (Davies, 1985). The cut-off distance or travel time, beyond which a consumer is unlikely to travel for a particular retail category (Ghosh, 1987) will be ascertained by primary data analysis.

Competition: The presence of existing shopping stores in a zone, called competition, directly and indirectly affects the market share of a store. This competitive interaction occurs at two major levels – economies of retail agglomeration (complementarity) and direct competition from analogous shops (Hotelling, 1929). Competition is observed to have a peculiar relationship with market share. It increases initially with a few analogous stores selling comparable goods and then diminishes when retail saturation is approached. The Tanner's modified gamma function is shown as $f(c_{ij}) = c_{ij}^{1.5} \exp(-2c_{ij})$ and has an asymmetric bell-shaped peak (Putman, 1983). It is assumed to reflect the dynamics of competition, since it closely corresponds to the identifiable retail life cycle with four distinct stages of retail competition – introduction, growth, maturity and decline (Dunne, 2002). Competition is measured by observed retail frontage of analogous and complementary trade and is an indication of the retail intensity and retail saturation of the location.

Urbanization: This component provides the attractiveness measure of the urban node itself, irrespective of what retail category location is being investigated. In effect, this component defines the ease of market entry for a shop (signifying location costs and disincentives) for each potential urban location. The four components used for calculating urbanization from a retail perspective are:

(1) Accessibility: availability of public transport, convenience in reaching, overall visibility
(2) Land-use: Urban restrictions, compatible landuse (commercial, mixed, public, recreational)
(3) Land availability & prices: Real estate prices, market trends, ease of market entry
(4) Market supporting Infrastructure: On-street parking, pedestrian aids, image

All the four components are measured on a psychometric scale – either on a five point scale from consumer and retailer surveys, or a three-point scale from expert surveys on urbanization.

3.2 The Proposed Model Formulation

If *FF* = market share (measured by observed footfall) and
LA = Location attractiveness and *SA* = Store attractiveness (input by order of floor area), then

$$FF = \frac{LA.SA}{\sum LA.SA}, \text{ and} \qquad \text{(Eq.1)}$$

$$LA = D^{\beta_1}.T^{-\beta_2}.C^{\beta_3^{1.5}\exp(-2\beta_3)}.U^{\beta_4} \qquad \text{(Eq.2)}$$

where, *D* = Demography, *T* = Travel time, *C* = competition and *U* = urbanization.

3.3 Calibration of the Model

The proposed Eq.2 is a special case of the Luce Choice axiom and may be generalized to the Multiplicative Competitive Interaction (MCI) form, as in

$$\pi_{ij} = \frac{(\prod_{k=1}^{q} x_{ij}^{\beta_k})\delta_{ij}}{[(\sum_{j=1}^{m}\prod_{k=1}^{q} x_{ij}^{\beta_k})\delta_{ij}]} \qquad \text{(Eq.3)}$$

where,

π_{ij} = probability that a customer in the *i*th choice situation selects the *j*th object

x_{kij} = the *k*th variable decribing object *j* in choice situation *i*

β_k = the parameter of sensitivity of π_{ij} with respect to variable *k*

δ_{ij} = disturbance term for model misspecification

The MCI formulation has desirable properties as a model of brand share determination, since the estimated market shares are guaranteed to be greater than zero and sum to unity, a condition which cannot be easily met by multivariate linear models (Naert & Bultez, 1973)

Eq.3 can be normally calibrated by Ordinary Least Squares (OLS) method as formulated by Nakanishi and Cooper (1974). This method of parameter estimation is employed in this study. The basic transformations used to 'balance' the model and normalize the primary data are briefly mentioned here. Eq.3 is transformed into a log-linear form with respect to the parameters of the model, the β_k's. Assume, $\log \delta_{ij}$ is independently and log-normally distributed with mean 0 and variance σ_δ^2. Let $\tilde{\pi}_i$, $\tilde{x}_{ki}$ and $\tilde{\delta}_i$ be the geometric means of π_{ij}, x_{ki} and δ_i over i, i.e., $\tilde{\pi}_i = (\prod_{j=1}^{m} \pi_{ij})^{1/m}$ Then, by taking logarithm of (Eq.3), substituting and rearranging terms, we get:

$$\log\left(\frac{\pi_{ij}}{\tilde{\pi}_i}\right) = \sum_{k=1}^{q} \beta_k \log\left(\frac{x_{kij}}{\tilde{x}_{ki}}\right) + \log\left(\frac{\delta_{ij}}{\tilde{\delta}_i}\right) \qquad \text{(Eq.4)}$$

Since Eq. 4 is linear in the β_k's, the use of multiple regression analysis is suggested to estimate the β_k's. Nakanishi and Cooper conclude that if sampling errors are negligible and specification errors are uncorrelated, simple log-linear transformations permits the use of OLS estimation programs, like simple regression models, as in

$$y_{ij} = \sum_{k=1}^{q} \beta_k z_{kij} + \varepsilon_{ij} \qquad \text{(Eq.5)}$$

For convenient application of data in the present study, (Eq.5) is represented in matrix notation.

$$y_i^{'} = (y_{i1} \quad y_{i2} \quad .... \quad y_{im})$$

$$Z_i = \begin{pmatrix} Z_{1i1} & Z_{2i1} & .... & Z_{qi1} \\ Z_{1i2} & Z_{2i2} & .... & Z_{qi2} \\ .... & .... & .... & .... \\ Z_{1im} & Z_{2im} & .... & Z_{qim} \end{pmatrix} \text{ and, } \quad y = \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_M \end{pmatrix}, \quad Z = \begin{pmatrix} Z_1 \\ Z_2 \\ ... \\ Z_M \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ ... \\ \varepsilon_M \end{pmatrix}$$

$$\varepsilon_i^{'} = (\varepsilon_{i1} \quad \varepsilon_{i2} \quad .... \quad \varepsilon_{im})$$

We thus, get the final log-linear transformed equation that is used for data input:

$y = Z\beta + \varepsilon$, where $\beta^{'} = (\beta_1, \beta_2, ..., \beta_q)$ (Eq.6)

3.4 The Interaction of Attractiveness measures – Location vs. Store

The study finally tries to assess the effectiveness of interaction between the two attractiveness measures used (*LA* and *SA*) for Location and Store. This is done by comparing three mathematical relationships – multiplication, addition and exponentiation – between the two measures. Case 1 is Eq.1 itself, $FF_1 = \frac{LA.SA}{\sum LA.SA}$. This is compared with 2 other formulations:

Case 2: $FF_2 = \frac{LA + SA}{\sum (LA + SA)}$, provided *SA>0* Case 3: $FF_3 = \frac{e^{LA} + e^{SA}}{\sum (e^{LA} + e^{SA})}$, provided *SA>0*

The condition *SA>0* implies the existence of a store, so that there can be footfall at the first place. It is apparent from the numerators, that Case 1 is a MCI formulation applied to the proposed model in general. Case 2 is a case of additive competitive interaction, similar to simple multiple regression with two independent variables. Finally, case 3 reflects the multinomial logit formulation, as widely used in mode choice studies in transportation planning.

The results obtained by using all three methods are tested and compared for the same set of data. Mean absolute percentage error (MAPE) estimates are used for comparison.

4. Primary Data Collection & Analysis

The city and retail category used for the case study have been carefully chosen. The target population considered in the study has been limited to the relevant socio-economic category likely to consume that particular product. Though wholesale commercial activities are beyond the scope of the choice domain, it may be noted that large front-end retailers themselves act as back-end wholesalers in the market supply chain.

4.1 Choice of case study: Organized clothing in Kolkata

Kolkata, formerly Calcutta, is one of India's oldest urban areas and historically the trading and commercial capital of India (till independence in 1947). It is India's second largest city with a

metropolitan area of about 1,900 sq. km and population of 16.3 million in 2005. It has a relatively high proportion of its population in the A & B socio-economic groups (35%) when compared with other Indian cities (JLL, 2007).

The urban Indian consumer as an aggregate has tremendous buying power, primarily because of its large size. India's purchasing power parity (PPP) is the fourth largest in the world and the per capita annual GDP (2003) was INR 23,408. Kolkata has a Market Potential value (MPV) of 613 (base Mumbai = 1000), the third highest after Mumbai (MPV base = 1000) and Delhi, but has relatively low Market Intensity Index (MII) value of 112, indicating greater growth opportunities (Marketing Whitebook, 2005). With its largely untapped customer base, a high and growing spending capacity and large skilled labour pool, Kolkata is presently a business-friendly city with a politically stable environment for corporates. These factors, coupled with the familiarity and proximity of the researcher to the city, dictated the choice of the case study city.

The convergence to the chosen case study – organized clothing stores selling readymade garments (excluding traditional Indian clothing) in the city of Kolkata is represented diagrammatically in Fig.5.

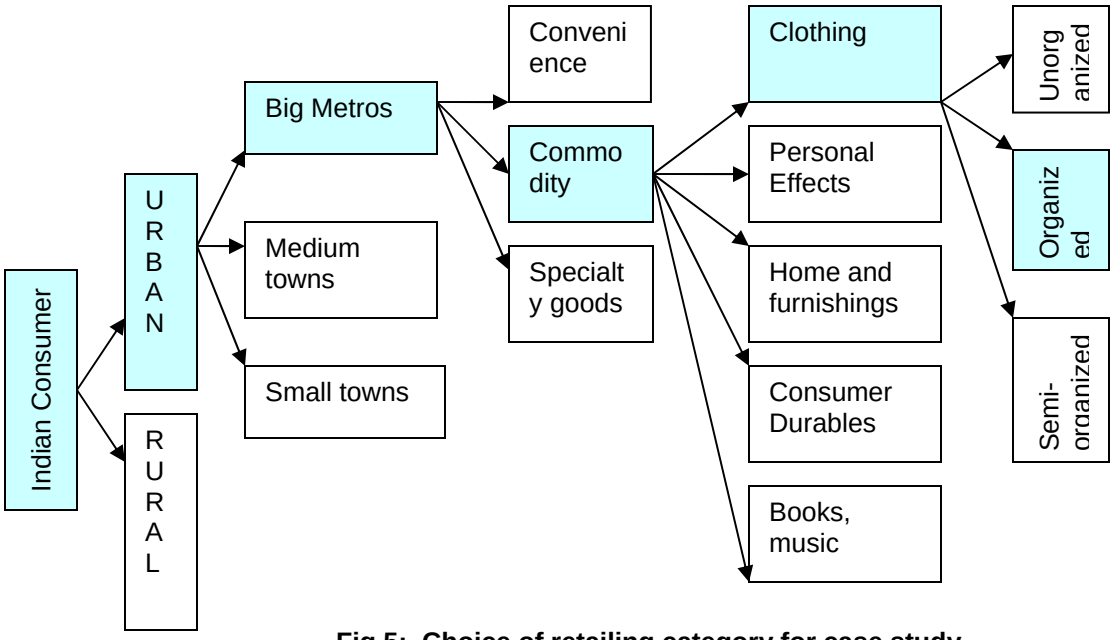


Fig 5: Choice of retailing category for case study

Market research indicates that convenience goods (food, groceries) are primarily travel time sensitive, since frequency of shopping is high. Specialty goods are increasingly chosen with the help of media advertisements and promotions. Commodity goods are thus most suited for the present study. Within this category, the average Indian urban consumers spends maximum of clothing and apparel – 10.6% of total expenditures in 2004 – and currently organized retailing predominantly targets to sell clothing to affluent Indian urbanites. Hence the chosen case study.

4.2 The survey procedures and analysis

Primary reconnaissance survey (2005) reveals the existing retail agglomerations for organized clothing stores of Kolkata. There are 14 such shopping zones, 8 of which are in inner-city locations and the other 6 in the outer-city locations (Fig.6). The existing shopping zones observed are:

(a) **Inner-city:** (1) Lindsay Street (2) Park Street - Chowringhee (3) Theatre Road – Camac Street (4) Elgin Road (5) Esplanade (6) Gariahat (7) Triangular Park (8) Hatibagan

(b) **Outer-City:** (9) VIP Rd (10) Salt Lake (11) E.M.Bypass (12) Chakgaria (13) Behala (14) Howrah.

In all these locations, all existing organized sector clothing stores are surveyed. For analysis, these 40 stores are categorized into three store formats, based on floor area of store:

(a) **Large format** (> 20,000 sqft.)	– 9 stores	(22.5% of sample)
(b) **Medium format** (5,000 - 20,000 sqft.)	– 13 stores	(32.5% of sample)
(c) **Small format** (< 5,000 sqft.)	– 18 stores	(45% of sample)

Consumer & Retailer surveys: Actual customers of these 40 shops were surveyed during peak festival season at peak shopping hours. Retailer surveys included questions regarding floor space, age of store and previous use of site, store format and positioning, branches and multiple stores, employment, footfall variation, customer conversion rates and average sales volume. The average footfall in each store is obtained from rigorous primary observations, for comparable time spans at comparable times of the day and week.

The total number of valid consumer surveys is 328, out of which 8 randomly selected observations are isolated and kept for later verification and model validation. Thus the primary customer survey dataset has 320 observations. About 75% of the observations were in inner-city locations and 25% in outer-city locations. The total number of customers surveyed in each shop varied from 6 to 12, based on their size and observed footfall. The survey instrument included questions regarding consumer demographic profile (age, gender, MHI), transport mode and willingness to travel, stated preferences for shopping zones and store formats, and also buying habits (frequency, repeat patronage, purpose of shopping trip and billing amount).

Each response is used for customer-spotting data (home and work origin vs. shopping destination) to calculate actual travel times. The consumers indicated the relative importance they attach to different store attributes (price, quality, range, service, interiors, infrastructure, advertising and store image) vis-à-vis different location attributes (public transportation availability, convenience in reaching store, visibility, ease of pedestrian movement, parking facilities, competition, complementarity and location image).

Primary data analysis is conducted to address four research questions, the findings of which are used later for model calibration

(a) Actual travel time vs. willingness to travel

(b) Changing consumer preferences for organized retail formats

(c) Relative importance of store attributes vs. location attributes

(d) Relative importance of shopping zone vs. individual store.

5.1 Data Application – normalized data inputs

Demographic, geographical and commercial data about the city is organized at this stage. The spatial distribution of the population, at the projected forecast time (2005), is assigned to the various urban nodes on the urban road network.

Demography: The spatial distribution of population of the intra-urban zones is obtained from census data along with reliable growth rate projections. The actual spending proportion of the resident population for the particular retail category is obtained from applied economic bureaus, while the socio-economic composition by monthly household income is obtained from newspaper readership surveys. Attributes such as number of target households, projected growth rate and density change are assigned to the nodes, as if the consumers reside on the node itself.

Road network: The digitized city map of Kolkata (NATMO, 2001) is assigned attributes for all the nodes and links – land use, land value and urban image. Speed data on roads along with node and link impedances (waiting time at road intersections) is used to calculate average travel time for all roads (links) in the network. The node-to-node travel time matrix is calculated by a computer code employing the Floyd's Algorithm (Taha, 1997). This widely used algorithm determines the shortest route between *any* two nodes in the network.

Retail Network: On the same digitized city map, the commercial and mixed land uses and existing retail locations are mapped, both for the commodity in study and other complementary retail types. This forms the retail database used to screen non-candidate locations and also to predict candidate new retail locations in the city.

5.2 Model validation & sensitivity analysis

The proposed model is validated by using Mean Absolute Percentage Error (MAPE) method of comparing predicted values with observed market share values, observed at a later time than the one previously used. It provided acceptable estimates of error.

Sensitivity analysis of the predicted attractiveness ranks is conducted across the three different store formats possible – large, medium and small – by changing the inputted aggregate values for store attractiveness (*SA)*. This provides insights on the suitability of locations for varying store size,

format and image. Finally, the interaction between the two attractiveness measures, LA and SA, are compared for three different formulations. It is established that MCI (multiplicative) is superior to the other two models of regression (additive) and multinomial logit (exponential) for measuring the competitive interaction of attractiveness determinants.

6. Findings and Conclusions

Based on the result outputs of the model application, as has been carefully interpreted in the fifth chapter of the final thesis, some notable findings and conclusions are mentioned here, under four related sub-headings: 1) prediction models 2) consumer behaviour 3) city 4) retail category.

6.1 Decision modelling related findings

The gravity model is a simple and effective strategy to predict relative market share of stores in intra-urban locations. The multiplicative competitive interaction formulation of the gravity model provides more tolerable estimates of market share than standard regression or multinomial logit formulations.

The two traditional parameters used in earlier gravity models (store area and travel time) are inadequate to accurately predict relative attractiveness of intra-urban nodes. A combination of 5 largely exclusive determinants of market share, viz. demography, travel time, competition, urbanization and store size, together provide a more reliable estimate of market share.

The log-linear transformation of the MCI model makes it possible to estimate the relative importance of the parameters, by using the OLS method. The coefficient values of the transformed regression form provide acceptable estimates of the exponential values to which the independent variables are raised. The coefficients generated for the proposed model in the case study are Demography = 0.862, Travel = - 1.447, Competition = 0.782 and Urbanization = 1.540. Their goodness of fit measures (R^2 values) and observed significance levels (p-values) are discussed in the thesis.

Among the two sets of 8 attributes of store and location choice, the ones with largest importance attached to them by consumers are Quality, Service and Product Range (for SA) and Access to public transport, Convenience in reaching store and Overall location image (for LA). Psychometric stated preferences show significant correlation between these attributes and observed market share. However different sets of attributes prove to be the better predictors (higher adjusted R^2 values) for different shopping zones.

The sensitivity analysis of the predicted attractiveness ranks for varying store formats indicate that location attractiveness and store attractiveness may be considered as independent and unrelated, i.e. any given intra-urban node commands a unique market share, that depends on its geographic position and urban setting, but not on the type or format of store located in it.

The prime reason for the ineffectiveness of prediction models in India is the unavailability of reliable time-series data on sales revenue. Responses to primary survey questionnaires reveal that retailers are generally unwilling to estimate or divulge accurate data on sales volume, customer conversion rates and total annual sales for the store for proprietary competitive reasons.

6.2 Consumer behaviour related findings

The age-cohort composition of the sample compared favourably with the age distribution of the overall city, thus proving the sample to be significant (5% confidence level) for drawing conclusions regarding consumer behaviour. The monthly household income (MHI) of the sample revealed that about 90% was in the socio-economic group-A, the target group initially assumed in the study. The majority of customers (48%) used public transport, while 9% walked for their shopping trips.

The average travel time for organized clothing in Kolkata is about 12.2 minutes. The traffic speed data and waiting time at intersections, roughly correspond 12 minutes travel time to a distance of about 5 kms. In reality this travel time may have increased values than the ones obtained from Traffic Police sources, especially in peak travel hours and peak shopping seasons, due to increased traffic congestion. Stated willingness for travel revealed higher values than observed travel times. The majority (40%) were willing to travel between 12 – 24 minutes, while 23% were willing to travel more than 24 minutes for annual festival shopping of clothing.

By stated preferences, there is a definite shift in consumer preferences over the last five years, to shop in large-store formats, distinctly establishing the rationale of the retail case study selection.

6.3 City related findings

The existing retail agglomerations for organized clothing in Kolkata indicates reveals two distinct higher order shopping zones located within the city – Lindsay Street and Gariahat – together accounting for more than two-third of the entire market share in 2005. A new mixed land-use zone, constituted by Theatre Road-Camac Street and Elgin Road, has emerged as the third-highest order shopping zone, accounting for 15% of the market share. The model predicts that these three zones will continue to enjoy the highest market shares followed by other new potential zones (Fig 6).

For better general comprehension of the predicted numerical market share values, three adjacent potential nodes may be grouped as an emerging shopping zone. The new emerging shopping zones, as predicted by the model, in order of market share ranks are: (1) Lake Gardens – Dhakuria – Jadavpur (2) Ultadanga – Kankurgachhi – Phoolbagan (3) Garia – Netaji Nagar – Baghajatin (4) E.M. Bypass – Science City – Sports Authority (5) Shyambazar – Bagbazar – Chidiya More (6) Taratola – Behala – Thakurpukur Bazar (7) PG-Elgin – Hazra – Rashbehari More (8) AJC Bose Road – Exide – Minto Park (9) Khidirpur – Ekbalpore – Bhawani Bhawan (10) Beadon Street – Girish Park – Vivekananda Road and (11) Bowbazar – Sealdah – Moulali.

EXISTING & PREDICTED
POTENTIAL SHOPPING ZONES IN KOLKATA

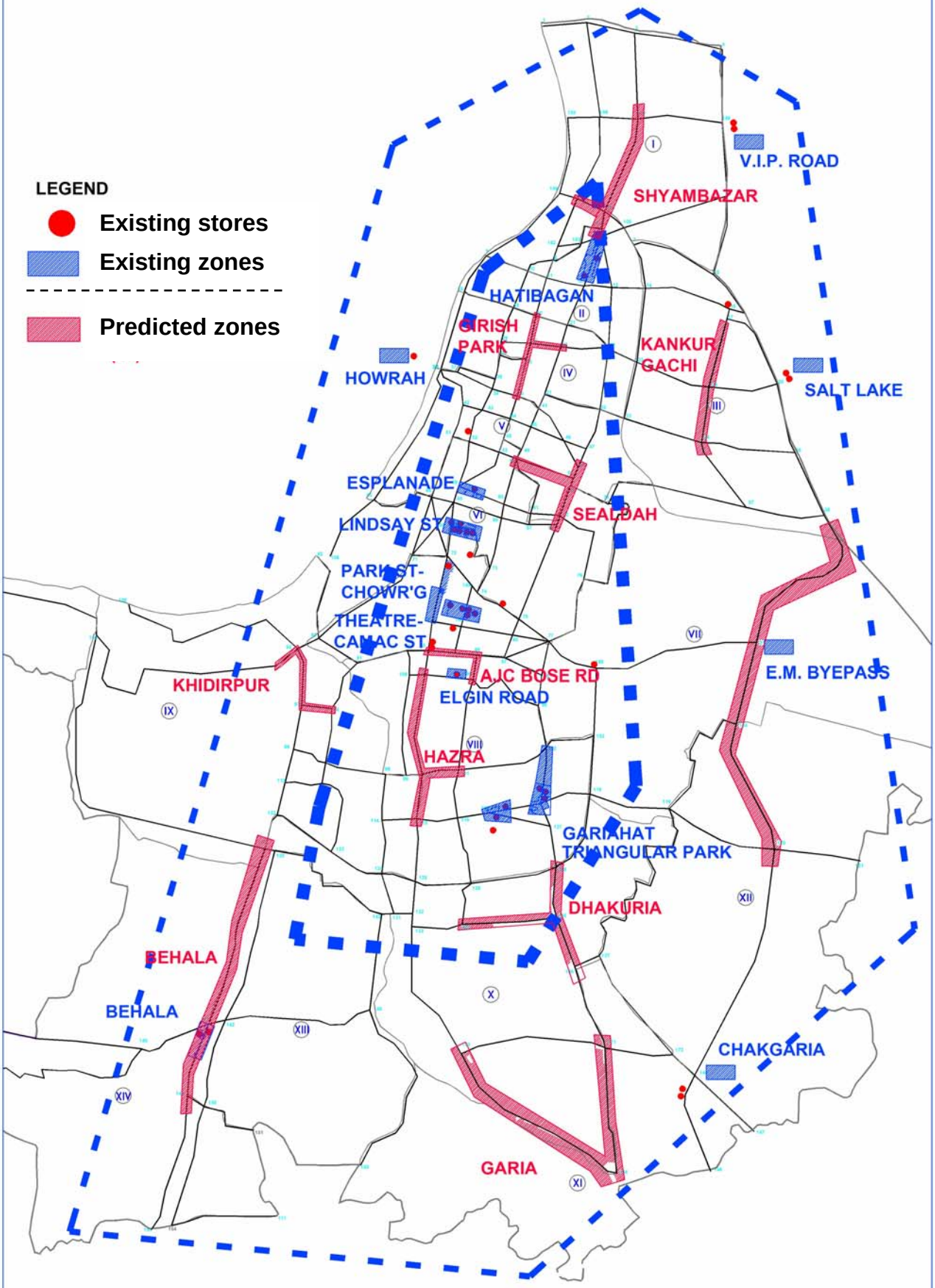


Fig.6: Existing & Predicted Potential Shopping Zones

The urban density of Kolkata Municipal area presently varies between 2500 to more than 10000 persons/sq.km. Between the time span of 1991 – 2001, the overall urban density has increased significantly in the outer portions of the city, while overcrowded central areas have reduced in density. Two major growth directions are apparent – a) southern fringes of Garia, Tollygunge, Behala and b) north-eastern fringe of E.M.Bypass, leading to Salt Lake, Lake Town and VIP Road.

The existing outer city retail locations are relatively recent developments and account for less than one-fifth of the entire market share for organized clothing retail. However, predicted values indicate larger future market shares for the outer-city zones. It is observed that most upcoming retail establishments in the city are coming up in conjunction with new residential and recreational projects, such as condominiums, malls and multiplexes. They are choosing sites with large land availability at relatively low prices and thus are most likely to locate in the outer city areas.

6.4 Organized Retail related findings

The total retail floor space under organized clothing in Kolkata (2005) was observed to be about 372,000 sqft, with an average floor space of 9,300 sqft per store. Format wise, the major share of floor space is under large formats (60%) and medium formats (25%). About 28% of the stores surveyed were newly opened, less than 2 years in operation, implying that organized retail sector is a relatively new phenomenon in Kolkata, presently at the growth stage of the competition cycle.

The majority of establishments surveyed (80%) were store multiples or chains, with atleast two outlets in the city (by 2005). The present bigger market players include Big Bazaar, Pantaloons, Westside and Shoppers' Stop, which serve as anchor stores for malls. The four operational shopping malls in Kolkata – Forum (Elgin Road), 22 Camac Street, City Centre (Salt Lake) and Metropolis Mall (Chakgaria) – are attracting many branded showrooms and non-anchor stores.

Organized retail trade for clothing provides significant urban employment in Kolkata, with 25% of the stores surveyed employing more than 100 full time employees, and about 40% of the stores employing between 21 and 50 people. The average floor space per employee is much lower for large-store formats (100-200 sqft/employee) than smaller formats (300-400 sqft/employee).

6.4 Discussions and Further Research

The decision model proposed in the present study is restricted in scope due to the non-availability of accurate and reliable data. Some areas of further research are mentioned here.

- The model may be generalized for any city and for any retail category
- The model may be developed into a user-friendly graphic user interface (GUI) for commercial use by prospective retailers. Then the present work will merit the status of a DSS.

- Geographic Information Systems (GIS) may be an ideal operating environment for such retail location models, with improved geometric analysis, display and presentation capabilities.
- The model may be formulated as a dynamic model, by adding a lagged variable term (E_{t-1}) to estimate the present potential (E_t) and also the rate of market share change.
- Analytical Hierarchical Process (AHP) may be employed to provide more reliable estimates of psychometric assessments for the sets of store and location attributes.
- Availability of reliable O-D data and travel mode-cost data may be used for research co-relating travel distance, travel time and generalized travel cost.
- From the point of view of urban planning studies, retail location studies provide an ideal platform for research in urban economics, combining transportation and land-use aspects. There is scope for further research in this multi-disciplinary domain, covering such aspects as commercial planning, land-use allocation, mixed land-use classification and urban infrastructure provision.

In conclusion, it is hoped that the proposed decision model developed by the researcher, will improve the quality of decision making in the field of retail location in Indian cities and encourage further research for better urban planning solutions in choice modelling and location allocation.

References

(for Synopsis only)

Arentze, T. and Timmermans, H. (2000), 'A Spatial DSS for Retail Plan Generation and Impact Assessment', *Transportation Research Part C: Emerging Technologies,* 8, 361-80.

ATKearney (2007), Growth Opportunities for Global Retailers: The 2007 Global Retail Development Index™. Chicago.

Biyani, K., CEO, Pantaloons India (2001), Interview in *Multimedia Series on Retailing.* Economic Times Intelligence Group.

Borgers, A. and Timmermans, H. (1991), 'A Decision Support and Expert System for Retail Planning', *Computers, Environment and Urban Systems,* 15, 179-88.

Census of India, 1991, 2001. Bureau of Applied Economics & Statistics (2003), *District Statistical Handbook: Kolkata*. Government of West Bengal, India.

CII: Confederation of Indian Industries (2004). *Retail Scenario in India: Unlimited opportunities.*

Carrothers, G. (1956), 'A historical review of the gravity and potential concepts of human interaction', *Journal of American Institute of Planners,* 22, 94-102.

Davies, R.L. and Rogers, D.S. eds. (1985), *Store Location and Store Assessment Research.* New York: John Wiley, 281-5.

Dunne, Lusch and Griffith (2002), *Retailing.* Fourth Ed. Singapore: South Western, Thomson.

Dutta J. and Banerjee U. (2002), 'Retail Location Theories and Prediction Techniques: An Overview of Various Approaches, 9(5), 18-27.

Fine, I.V. (1954), 'Retail trade area analysis: A guide to effective use of retail trade area studies', *Wisconsin Commerce Papers*, 1 (6).

Ghosh, A and McLafferty, S. (1987), *Location Strategies for Retail and Service Firms.* Lexington, Massachusetts: Lexington Books.

Gilbert, D. (1999), *Retail Marketing Management.* Essex: Pearson Education Limited, 228.

Harris, B. (1964), 'A Note on the probability of interaction at a distance', *Journal of Regional Science,* 5, 31-5.

Hotelling, H. (1929), 'Stability in Competition', *Economic Journal,* 34, 41-57.

Huff, D.L. (1963), 'A Probabilistic Analysis of Shopping Centre Trade Areas', *Land Economics,* 39, 81-90.

IMRB: Indian Market Research Bureau (2001), Indian Readership Survey. National Readership Survey.

Jones Long LaSalle (2007), *India – Emerging City Winners: Kolkata.*

Keen and Scott-Morton (1978), *Decision Support Systems: An Organizational Perspective.* Massachusetts: Addison Wesley.

Kivell, P.T. and Shaw, G. (1980), 'The Study of Retail Location', in *Retail Geography* (ed. Dawson, J.A. London: Croom Helm.

Lakshmanan and Hansen (1965), 'A Retail Market Potential Model', *Journal of American Institute of Planners,* 31, 134-44.

Lösch, A. (1954), *The Economics of Location.* New Haven: Yale University Press.

Lowry, I.S. (1964), *Model of the Metropolis*. Rand Corporation Memo, RM-4035-RC.

Marketing Whitebook (2005), Business World. Kolkata: ABP Publication.

Mejia, L.C. and Benjamin, J.D. (2002), 'What do we know about the determinants of shopping centre sales? Spatial vs. non-spatial factors', *Journal of Real Estate Literature.*

Naert, P.A. and Bultez, A.V. (1973), 'Logically Consistent Market Share Models,' *Journal of Marketing Research,* 10, 334-40.

Nakanishi, M. and Cooper, L.G. (1974), 'Parameter Estimation for a Multiplicative Competitive Interaction Model – Least Squares Approach', *Journal of Marketing Research,* 11, 303-11.

NATMO: National Atlas Thematic Mapping Organization (2001), *District Planning Map Series: Kolkata*

NCAER (National Council for Applied Economic Research) (2001), *Market Information Survey of Indian Households.*

VISION (1987), a DSS developed by National Decision Systems (1987), USA.

ORG MARG (1995), Census of Retail Establishments, India (for FMCG products).

Putman, S.H. (1983), *Integrated Urban Models – Policy Analysis of Transportation and Land Use.* London: Pion Limited.

Reilley, W.J. (1931), *The Laws of Retail Gravitation.* New York: Knickerbocker Press.

ReVelle, Marks and Liebman (1970), 'An analysis of Public and Private Sector Location Models', *Management Science,* 16, 692-707.

Satani, N. et al (1998), 'Commercial Facility Location Model using Multiple Regression Analysis', *Computers, Environment & Urban Systems,* 22(3), 219-40.

Scott, P. (1970), *Geography and Retailing.* London: Hutchison University Press.

Sprague and Carlson (1974), 'Bit by bit: Towards Decision Support Systems', *California Management Review*, 22, 60-7.

Taha, H.A. (1997), *Operations Research – An Introduction.* Sixth Ed. New Jersey: Prentice Hall.

Turban, E. (1990), *Decision Support and Expert Systems: Management support systems.* New York: Macmillan.

Wilson, A.J. (1971), 'A Family of Spatial Interaction Models and associated development', *Environment and Planning,* 3, 1-32.